ON ASYMPTOTIC PROPERTIES OF THE LEAST SQUARES ESTIMATORS IN A STATIONARY RANDOM FIELD

S. Nandi and D. Kundu

Technical Report 99-07

July 1999

Center for Multivariate Analysis
417 Thomas Building
Penn State University
University Park, PA 16802

r

19991101 081

REPORT DOCUMENTATION PAGE		Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE July 1999	3. REPORT TYPE AND DATES COVERED Technical - 99-07	

4. TITLE AND SUBTITLE On Asymptotic Properties of the Least Squares Estimators in a Stationary Random Field	5. FUNDING NUMBERS DAAH04-96-1-0082
6. AUTHOR(S) S. Nandi and D. Kundu	

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES) Center for Multivariate Analysis 417 Thomas Building Department of Statistics Penn State University University Park, PA 16802	8. PERFORMING ORGANIZATION REPORT NUMBER 99-07

9. SPONSORING / MONITORING AGENCY NAMES AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING / MONITORING AGENCY REPORT NUMBER ARO 35518.53-MA

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release: distribution unlimited.	12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

ABSTRACT: A particular two dimensional model in a stationary random field, which has a wide applications in statistical signal processing and in texture classifications, is considered. We prove the consistency and also obtain the asymptotic distributions of the least squares estimators of the different model parameters. It is observed that the asymptotic distribution of the least squares estimators are multivariate normal. Some numerical experiments are performed to see how the asymptotic results work for finite samples. We propose some open problems at the end.

14. SUBJECT TERMS Strong consistency, Texture Classification, Statistical Signal Processing, Stationary random field	15. NUMBER IF PAGES 10
	16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

ON ASYMPTOTIC PROPERTIES OF THE LEAST SQUARES ESTIMATORS IN A STATIONARY RANDOM FIELD.

Swagata Nandi and Debasis Kundu
Department of Mathematics
Indian Institute of Technology Kanpur
Kanpur, Pin 208016 INDIA
nandi@iitk.ac.in
kundu@iitk.ac.in

ABSTRACT: A particular two dimensional model in a stationary random field, which has a wide applications in statistical signal processing and in texture classifications, is considered. We prove the consistency and also obtain the asymptotic distributions of the least squares estimators of the different model parameters. It is observed that the asymptotic distribution of the least squares estimators are multivariate normal. Some numerical experiments are performed to see how the asymptotic results work for finite samples. We propose some open problems at the end.

AMS Subject Classification: $62F12, 62M10$.

Key words: Strong consistency, Texture Classification, Statistical Signal Processing, Stationary random field.

Short Running Title: Two Dimensional Model.

Corresponding Author: Swagata Nandi

1. INTRODUCTION:

We consider the following two dimensional model:

$$y(m,n) = \sum_{k=1}^{q} A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) + X(m,n);$$

$$\text{for } m = 1, \ldots M, n = 1, \ldots, N, \qquad (1.1)$$

where A_k^0's are unknown real numbers, λ_k^0 's, μ_k^0 's are unknown frequencies. For identifiability, we need to assume $\lambda_k^0 \in (-\pi, \pi)$ and $\mu_k^0 \in (0, \pi)$ and they are distinct. $X(m,n)$ is a two dimensional (2-D) stationary random field described as follows:

$$X(m,n) = \sum_{i=-P}^{P} \sum_{j=-Q}^{Q} b(i,j)e(m-i, n-j). \qquad (1.2)$$

Here $\{e(m,n)\}$ is a two dimensional sequence of independent and identically distributed ($i.i.d.$) random variable with mean zero and finite variance. P and Q are arbitrary positive integers. 'q', the number of components, is assumed to be a known integer. Given a sample $y(m,n); m = 1, \ldots, M, n = 1, \ldots, N$, the problem is to estimate A_k^0 's, λ_k^0 's, μ_k^0's for $k = 1, \ldots q$.

Here $X(m,n)$ is a stationary random field and $y(m,n)$ is a non-stationary random field. To see how this model represents different textures, the readers are referred to the work of Mandrekar and Zhang [17] or Francos *et al.* [2], where they provided nice 2-D image plots of $y(m,n)$, whose grey level at (m,n) is proportional to the value of $y(m,n)$ and when it is corrupted by independent Gaussian noise field. So this model represents mixed textures of regular textures with noise pictures. Our problem is to extract the regular textures from the contaminated $y(m,n)$. The problem is of interest in spectrographs and is studied using group theoretic methods by Malliavan [15,16]. Francos *et al.* [2] considered the Wold type decomposition of the random fields due to Helson and Lowdenslager [5,6] but no concrete mathematical results were obtained in that paper. Mandrekar and Zhang [17] also considered the spectral analysis of this problem under the following stationary assumptions on $X(m,n)$

$$X(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} b(i,j)e(m-i, n-j), \qquad (1.3)$$

1

where $\{e(m,n)\}$ is a double array sequence of independent random variables such that

$$\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |b(i,j)| < \infty,$$

$$E(e(m,n)) = 0, \ E(|e(m,n)|^r) < \infty, \qquad (1.4)$$

for some constant $r > 2$. They proved that the spectral estimators of λ's and μ's are consistent estimators of the corresponding parameters when $X(m,n)$ satisfies (1.3) and (1.4). Unfortunately the corresponding estimators of the linear parameters (A's) are not consistent. Moreover, they could not obtain the asymptotic distribution of the different estimators. Therefore, the rates of convergence of those estimators are not known. Their results are mainly based on the work of Lai and Wei [14], which is quite involved mathematically. In this paper we mainly consider the least squares estimators (LSE's) of the different parameters and study their large sample properties.

It may be argued that the assumption of Mandrekar and Zhang [17] on $X(m,n)$ is somewhat weaker than ours, because in our case $P < \infty$ and $Q < \infty$ as defined in (1.2). But since P and Q are arbitrary, therefore (1.3) can be approximated arbitrary closely by (1.2) with sufficiently large P and Q, see Fuller [3]. Therefore, for all practical purposes they are equivalent. More over Mandrekar and Zhang [17] use higher order moment assumptions ($r > 2$) on $e(m,n)$ to prove the necessary consistency results, whereas we assume only the finite second moment of $e(m,n)$, to prove the consistency and the asymptotic normality of the LSE's of all the unknown parameters. In this paper the almost sure convergence means with respect to the usual Lebesgue measure and it will be denoted by a.s. We will denote the set of positive integers by Z. Also the notation $a = O(b(M,N))$, means, $|a/b(M,N)|$ is bounded for all M and N.

The rest of the paper is organized as follows. In Section 2, we prove the strong consistency and in Section 3 we obtain the asymptotic distributions of the LSE's of the parameters of the model (1.1), when $q = 1$. For $q > 1$, the results are obtained in Section 4. We perform some numerical experiments and present those results in Section 5 and finally we draw conclusions from our work and propose some open problems in Section 6.

2. CONSISTENCY OF THE LSE'S:

In this section, we obtain the consistency of the LSE's of the unknown parameters of the model (1.1), when $q = 1$, i.e.,

$$y(m,n) = A^0 cos(m\lambda^0 + n\mu^0) + X(m,n);$$

$$\text{for } m = 1, \ldots M, n = 1, \ldots, N.$$

The LSE's are obtained by minimizing $Q(\theta)$, where

$$Q(\theta) = \sum_{m=1}^{M} \sum_{n=1}^{N} (y(m,n) - Acos(m\lambda + n\mu))^2. \quad (2.1)$$

Here $\theta = (A, \lambda, \mu)$, the true parameter value and the LSE of θ are denoted by $\theta^0 = (A^0, \lambda^0, \mu^0)$ and $\hat{\theta} = (\hat{A}, \hat{\lambda}, \hat{\mu})$ respectively. We make it explicit the assumptions on $X(m,n)$ as follows.

Assumption 1: Let $\{X(m,n) \in Z\}$ be a stationary random field and each $X(m,n)$ can be represented as (1.2), $\{e(m,n) \in Z\}$ is a double array sequence of *i.i.d.* random variables with mean zero and variance σ^2.

We use the following lemma to prove the necessary results.

Lemma 1 If the double array sequence $\{X(m,n); m,n \in Z\}$ satisfy Assumption 1, then

$$\sup_{\alpha,\beta} |\frac{1}{N} \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n)cos(m\alpha)cos(n\beta)| \overset{a.s.}{\to} 0$$

$$\text{when } \min\{M,N\} \to \infty.$$

Proof: See Appendix.

Note that the Lemma 1, is a very strong result. It extends some of the existing one dimensional results of Hannan [4], Walker [25], Rao and Zhao [22], Kundu [8] and Kundu and Mitra [12,13] to the 2-D case. It also generalizes the multidimensional results of Bai et al. [1], Rao et al. [23], Kundu and Mitra [11] and Kundu and Gupta [10] in some sense.

Consider the following assumption on the parameters of the model (1.1), when $q = 1$.

Assumption 2: Let A^0 be a arbitrary real number not identically equal to zero, $\lambda^0 \in (-\pi, \pi)$ and $\mu^0 \in (0, \pi)$.

Now we state the consistency result as the following theorem.

Theorem 1: Under the assumptions 1 and 2, the least squares estimators of the parameters of the model (1.1) when $q = 1$, are strongly consistent.

Proof of Theorem 1: Expanding (2.1), with the help of Lemma 1 and using the similar technique as of Bai *et al.* [1], the results can be obtained.

It is interesting to observe that although the errors are correlated the usual LSE's provide consistent solutions. For the general linear or non-linear models if the errors are correlated, it is well known (Rao; [20], Seber and Wild; [24]) that the usual LSE's are inconsistent. In the correlated case, we need to consider the generalized least squares estimators, which are consistent. On the other hand, theorem 1, may not be too surprising, because it is known (Kundu; [9]) that for one-dimensional frequency model, even if the errors are correlated, the LSE's are consistent. In this respect one or higher dimensional frequency models are quite different than the usual non-linear models.

3. ASYMPTOTIC NORMALITY OF THE LSE'S:

In this section we obtain the asymptotic distributions of the least squares estimators of the parameters of the model (1.1) when $q = 1$. We use the following notations. The first derivative of $Q(\theta)$ is a 1×3 vector as

$$Q'(\theta) = \left[\frac{\delta Q(\theta)}{\delta A}, \frac{\delta Q(\theta)}{\delta \lambda}, \frac{\delta Q(\theta)}{\delta \mu} \right]$$

and the second derivative is a 3×3, matrix as follows;

$$Q''(\theta) = \begin{bmatrix} \frac{\delta^2 Q(\theta)}{\delta A^2} & \frac{\delta^2 Q(\theta)}{\delta A \delta \lambda} & \frac{\delta^2 Q(\theta)}{\delta A \delta \mu} \\ \frac{\delta^2 Q(\theta)}{\delta \lambda \delta A} & \frac{\delta^2 Q(\theta)}{\delta \lambda^2} & \frac{\delta^2 Q(\theta)}{\delta \lambda \delta \mu} \\ \frac{\delta^2 Q(\theta)}{\delta \mu \delta A} & \frac{\delta^2 Q(\theta)}{\delta \mu \delta \lambda} & \frac{\delta^2 Q(\theta)}{\delta \mu^2} \end{bmatrix}.$$

Therefore expanding $Q'(\hat{\theta})$ around θ^0, we obtain

$$Q'(\hat{\theta}) - Q'(\theta^0) = (\hat{\theta} - \theta^0) Q''(\bar{\theta}) \qquad (3.1)$$

where $\bar{\theta}$ is a point on the line joining between the points $\hat{\theta}$ and θ^0. Note that $Q'(\hat{\theta}) = 0$ and consider the 3×3 diagonal matrix $\mathbf{D}$ as follows.

$$\mathbf{D} = \begin{bmatrix} M^{-\frac{1}{2}} N^{-\frac{1}{2}} & 0 & 0 \\ 0 & M^{-\frac{3}{2}} N^{-\frac{1}{2}} & 0 \\ 0 & 0 & M^{-\frac{1}{2}} N^{-\frac{3}{2}} \end{bmatrix}.$$

Now (3.1) can be written as

$$(\hat{\theta} - \theta^0) = -Q'(\theta) \left[Q''(\bar{\theta}) \right]^{-1} \qquad (3.2)$$

if $Q''(\bar{\theta})$ is a full rank matrix (see at the end of this section). Equivalently

$$(\hat{\theta} - \theta^0) \mathbf{D}^{-1} = -[Q'(\theta^0) \mathbf{D}][\mathbf{D} Q''(\bar{\theta}) \mathbf{D}]^{-1}. \qquad (3.3)$$

Now let's consider different elements of $\left[Q'(\theta^0) \mathbf{D} \right]$.

$$\frac{1}{M^{\frac{1}{2}} N^{\frac{1}{2}}} \frac{\delta Q(\theta^0)}{\delta A} = \frac{-2}{M^{\frac{1}{2}} N^{\frac{1}{2}}} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n) \\ cos(m\lambda + n\mu),$$

$$\frac{1}{M^{\frac{3}{2}} N^{\frac{1}{2}}} \frac{\delta Q(\theta^0)}{\delta \lambda} = \frac{2}{M^{\frac{3}{2}} N^{\frac{1}{2}}} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n) \\ Amsin(m\lambda + n\mu),$$

$$\frac{1}{M^{\frac{1}{2}} N^{\frac{3}{2}}} \frac{\delta Q(\theta)}{\delta \mu} = \frac{2}{M^{\frac{1}{2}} N^{\frac{3}{2}}} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n) \\ Ansin(m\lambda + n\mu).$$

Using the central limit theorem of the stochastic process (see Fuller; [3]), and using the following results of Mangulis [18] for $\beta \neq 0$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} cos^2(t\beta) = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} sin^2(t\beta) = \frac{1}{2}$$

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t cos^2(t\beta) = \lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t sin^2(t\beta) = \frac{1}{4}$$

$$\lim_{n \to \infty} \frac{1}{n^3} \sum_{t=1}^{n} t^2 cos^2(t\beta) = \lim_{n \to \infty} \frac{1}{n^3} \sum_{t=1}^{n} t^2 sin^2(t\beta) = \frac{1}{6}$$

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t sin(t\beta) cos(t\beta) = 0$$

3

it follows that $[Q'(\theta^0)\mathbf{D}]$ tends to a 3-variate normal distribution with mean vector zero and the dispersion matrix $2\sigma^2 c\Sigma$, where

$$c = \left|\sum_{i=-P}^{P}\sum_{j=-Q}^{Q} b(i,j)cos(i\lambda^0)cos(j\mu^0)\right|^2 + \quad (1)$$

$$\left|\sum_{i=-P}^{P}\sum_{j=-Q}^{Q} b(i,j)cos(i\lambda^0)sin(j\mu^0)\right|^2 +$$

$$\left|\sum_{i=-P}^{P}\sum_{j=-Q}^{Q} b(i,j)sin(i\lambda^0)cos(j\mu^0)\right|^2 + \quad (2)$$

$$\left|\sum_{i=-P}^{P}\sum_{j=-Q}^{Q} b(i,j)sin(i\lambda^0)sin(j\mu^0)\right|^2 \quad (3.4)$$

and

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3}A^{0^2} & \frac{1}{4}A^{0^2} \\ 0 & \frac{1}{4}A^{0^2} & \frac{1}{3}A^{0^2} \end{bmatrix}. \quad (3.5)$$

Observe that because of theorem 1, $\bar{\theta}$ converges to θ^0 a.s. and

$$\lim_{M,N\to\infty} (\mathbf{D}Q''(\bar{\theta})\mathbf{D}) = \lim_{M,N\to\infty} (\mathbf{D}Q''(\theta^0)\mathbf{D}) = \Sigma. \quad (3.6)$$

Therefore from (3.3), we have the following result

Theorem 2: Under the assumptions 1 and 2, the limiting distribution of $\{M^{\frac{1}{2}}N^{\frac{1}{2}}(\hat{A}-A^0), M^{\frac{3}{2}}N^{\frac{1}{2}}(\hat{\lambda}-\lambda^0), M^{\frac{1}{2}}N^{\frac{3}{2}}(\hat{\mu}-\mu^0)\}$ as Min(M,N) $\to \infty$, is a 3-variate normal with mean vector zero and covariance matrix $2\sigma^2 c\Sigma^{-1}$, when Σ^{-1} has the following structure:

$$\Sigma^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{48}{7}\frac{1}{A^{0^2}} & -\frac{36}{7}\frac{1}{A^{0^2}} \\ 0 & -\frac{36}{7}\frac{1}{A^{0^2}} & \frac{48}{7}\frac{1}{A^{0^2}} \end{bmatrix}. \quad (3.7)$$

Note that to prove theorem 2, we use the fact $Q'(\bar{\theta})$ is a full rank matrix a.s. for large M and N. In fact, we have used $\mathbf{D}Q''(\bar{\theta})\mathbf{D}$ is of full rank a.s. (see (3.3)). Now from (3.6), it is clear that for large M and N, $\mathbf{D}Q''(\theta^0)\mathbf{D}$ is a full rank matrix. Since, the elements of the matrix $Q''(\theta)$ are continuous functions of θ and $\bar{\theta}$

converges to θ^0 a.s., therefore $\mathbf{D}Q''(\bar{\theta})\mathbf{D}$ is a full rank matrix a.s. for large M and N.

From theorem 2, it is clear that the LSE of the amplitude (A's) is asymptotically independent with the LSE's of the frequencies. Where as, the LSE's of the two frequencies have a high negative correlation. The asymptotic variances of the LSE's of A, λ and μ are proportional to $\frac{1}{MN}$, $\frac{1}{M^3 N A^{0^2}}$ and $\frac{1}{M N^3 A^{0^2}}$ respectively. Therefore, it is immediate that the convergence rates of $\hat{\lambda}$ and $\hat{\mu}$ are of the orders $O(M^{-3}N^{-1})$ and $O(M^{-1}N^{-3})$ respectively and both of them are faster than the convergence rate of $\hat{A}$, which is $O((MN)^{-1})$. More over, the asymptotic variances of $\hat{\lambda}$ and $\hat{\mu}$ are inversely proportional to A^{0^2}. This may not be very surprising, because if A^{0^2} is small, then it is difficult to estimate the frequencies.

4. MULTIPARAMETER CASE:

In this section we consider the model (1.1) for any integer q. We use the following notations

$$\theta_1 = (A_1, \lambda_1, \mu_1), \ldots, \theta_q = (A_q, \lambda_q, \mu_q),$$

$$\Psi = (\theta_1, \ldots, \theta_q)$$

The true parameter value and the LSE's of Ψ will be denoted by Ψ^0 and $\hat{\Psi}$ respectively. We investigate the consistency and the asymptotic properties of $\hat{\Psi}$, which is obtained by minimizing

$$R(\Psi) = \sum_{m=1}^{M}\sum_{n=1}^{N} \left(y(m,n) - \sum_{k=1}^{q} A_k cos(m\lambda_k + n\mu_k)\right)^2$$

with respect to Ψ. We need the following **assumption**.

Assumption 3: Let $A_1^0, \ldots, A_p^0$ be arbitrary real numbers not any one of them are identically equal to zero, $\lambda_1^0, \ldots, \lambda_q^0 \in (-\pi, \pi)$ and they are distinct, similarly $\mu_1^0, \ldots, \mu_q^0 \in (0, \pi)$ and they are distinct.

The following result provides the consistency results of the LSE's of the model parameter for the general case.

Theorem 3: Under assumptions 1 and 3, $\hat{\Psi}$ is a strongly consistent estimator of Ψ^0.

4

Proof: It is quite similar to the proof of Theorem 1, so it is omitted.

To establish the asymptotic distribution of $\hat{\Psi}$, we use the following notations. The $3q \times 3q$ diagonal matrix $\mathbf{V}$ and the $3q \times 3q$ block diagonal matrix Φ^{-1} are defined as follows.

$$\mathbf{V} = \begin{bmatrix} \mathbf{D} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{D} \end{bmatrix},$$

$$\Phi^{-1} = \begin{bmatrix} c_1 \Sigma_1^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & c_q \Sigma_q^{-1} \end{bmatrix}$$

where Σ_j^{-1} can be obtained from Σ^{-1} defined in (3.6) by replacing A^{0^2} with $A_j^{0^2}$ and similarly c_k's can be obtained from c's defined in (3.4) by replacing λ^0 and μ^0 with λ_k^0 and μ_k^0 respectively.

Theorem 4: Under the same assumptions as Theorem 3, $(\hat{\Psi} - \Psi^0)\mathbf{V}^{-1}$ converges to a $3q$-variate normal distribution with mean vector zero and the dispersion matrix $2\sigma^2 \Phi^{-1}$, where $\mathbf{V}^{-1}$ and Φ^{-1} are as defined above.

Proof: The proof can be obtained quite similarly as Theorem 2, so it is omitted.

5. NUMERICAL EXPERIMENTS AND DISCUSSIONS:

In this section we present some results of the numerical experiments performed to see how the asymptotic results behave for finite sample sizes. We performed all the experiments in Silicon Graphics, using the random deviate generator of Press et al. [19]. We considered the following model:

$$y(m, n) = 4.0cos(2.0m + 1.0n) + 5.0cos(2.5m + 1.5n)$$

$$+ X(m, n), \qquad (5.1)$$

$X(m, n)$ has the following form

$$X(m, n) = e(m, n) + .25e(m - 1, n) + .25e(m + 1, n)$$

$$+ .25e(m, n - 1) + .25e(m, n + 1)$$

$\{e(m, n); m = 1, \ldots, M, n = 1, \ldots, N\}$ are *i.i.d.* Gaussian random variables with mean zero and finite variance σ^2. The stationary random field $X(m, n)$ has that particular structure indicates that the error at the point (m, n) is equally influenced by the four equidistant points from (m, n). We considered $M = N = 10, 20, 30, 40, 50$ and $\sigma = .25, .50, .75, 1.0$. For each sample size and for each σ we computed the LSE's of $A_1, A_2, \lambda_1, \lambda_2, \mu_1$ and μ_2 and observed the average estimates and the average mean squared errors (MSE's) over five hundred replications. We present the results in Tables 1-5. We also report the asymptotic variances (ASV) for each parameters for comparison purposes.

From the simulations it became very clear that as sample size increases or the variance decreases, the average MSE's and biases of all the estimators decrease. It shows that all the estimators are consistent and asymptotically unbiased. Biases are quite small even when the sample sizes are quite small. It is clear that the MSE's of the estimators of the non-linear parameters are smaller than that of the linear parameters even for small sample sizes. From the experimental study also it is clear that the estimation of the linear parameters are more difficult (in terms of accuracy) compared to the non-linear parameters. Some of the asymptotic behaviors are present even at small sample sizes. For example if $A_1 < A_2$, then it is observed that the MSE's of $\hat{\mu}_2$ and $\hat{\lambda}_2$ are smaller than that of $\hat{\mu}_1$ and $\hat{\lambda}_1$ respectively. It is also observed that as sample size increases the MSE's become closer to the asymptotic variances, *i.e.* |ASV - MSE | decreases. Therefore looking at the behavior of the MSE's we can say that the asymptotic results can be used to draw the small sample inferences for the different model parameters. In some of the cases it is observed that the ASV is lower than the corresponding MSE. This may not be very surprising, since we considered only five hundred replications, it may be due to the sampling errors (see Karian and Dudewicz; [7]).

6. CONCLUSIONS:

In this paper we consider the estimation of the parameters of a two dimensional model, which has wide applicability in Statistical Signal Processing and in Tex-

ture classifications. We study the asymptotic properties of the LSE's of the model parameters and show that the LSE's are strongly consistent. We also obtain the asymptotic distributions of the LSE's, which provides the rate of convergence of the LSE's. This paper generalizes some of the existing one dimensional results to the 2-D case. It generalizes some of the multidimensional results also in certain way. Numerical experiments suggest that the asymptotic results can be used to draw the small sample inferences for the linear and non-linear parameters. We do not address one important problem, namely the estimation of 'q'. That is a very important problem in practice. We may have to use certain information theoretic criteria like AIC, BIC or we may have to use the cross validation type technique as proposed by Rao [21] for the one-dimensional case. Another important problem is to obtain an efficient estimator of the different parameters by some non-iterative technique. Non iterative techniques are important for online implementations or to use as initial guesses for any iterative procedures. More work is needed in these directions.

REFERENCES:

[1] Bai, Z.D., Kundu, D. and Mitra, A. (1996), "A theorem in probability and its applications in multidimensional signal processing", *IEEE Trans. Signal Processing*, Vol. 44, No. 12, pp. 3167-3169.

[2] Francos, J.M., Mieri, A.Z. and Porat, B. (1990) "On a Wold type decomposition of 2-D discrete random field", *Proc. Int. Conf. Acoust. Speech. and Signal Processing*, Albuquerque.

[3] Fuller, W. (1976), *Intriduction to Statistical Time Series*, John Wiley and Sons, New York.

[4] Hannan, E.J. (1971) "Nonlinear Time Series regression", *Jour. Applied Probability*, Vol. 8, pp. 767-780.

[5] Helson, H. and Lowdenslager, D. (1958), "Prediction theory and Fourier series in several variables", *Acta Math.*, Vol. 99 pp. 165-202.

[6] Helson, H. and Lowdenslager, D. (1962), "Prediction theory and Fourier series in several variables", *Acta Math.*, Vol. 106, pp. 175-213.

[7] Karian, Z.A. and Dudewicz, E.J. (1999), *Modern Statistical Systems and GPSS Simulations*, 2nd. Edition, CRC Press, Washington, D.C.

[8] Kundu, D. (1993), "Asymptotic theory of the least squares estimator of a particular non-linear regression model", *Statistics and Probability Letters*, Vol. 18, pp. 13-17.

[9] Kundu, D. (1997), "Asymptotic theory of the least squares estimators of sinusoidal signal", *Statistics*, Vol. 30, pp. 221-238.

[10] Kundu, D. and Gupta, R.D. (1998), "Asymptotic properties of the least squares estimators of a two dimensional texture model", *Metrika*, Vol. 48, pp. 83-97.

[11] Kundu, D. and Mitra, A. (1996a), "Asymptotic properties of the least squares estimators of 2-D exponential signals", *Multidimensional Systems and Signal Processing*, Vol. 7, No. 3, pp. 135-150.

[12] Kundu, D. and Mitra, A. (1996b), "Asymptotic theory of the least squares estimator of a non-linear regression model", *Communications in Statistics, Theory and Methods*, Vol. 25, No. 1, pp. 133-141.

[13] Kundu, D. and Mitra, A. (1999), "On asymptotic properties and confidence intervals of superimposed exponential signals", *Signal Processing*, Vol. 22,

6

no. 2, pp. 129-139.

[14] Lai, T.L. and Wai, Z.C. (1982) "A law of iterated logarithm for double array of independent random variables with applications to regression and time series models", *Ann Probability*, Vol. 10, No. 2, pp. 320-335.

[15] Malliavin, P. (1994a), "Sur la norme d'une matrice circulante gaussienne", *C.R. Acad. Sci. Paris t. 319, Serie I*, pp. 745-749.

[16] Malliavin, P. (1994b), "Estimation d'un signal lorentzien", *C.R. Acad. Sci. Paris t. 319, Serie I*, pp. 991-997.

[17] Mandrekar, V. and Zhang H. (1995), "Determination of discrete spectrum of random fields with application to texture classification", Technical Report, Michigan State University, U.S.A.

[18] Mangulis, V. (1965), *Handbook of Series for Scientists and Engineers*, Academic Press, New York.

[19] Press et al. (1993) *Numerical recipes in Fortran*, Cambridge University Press, Cambridge.

[20] Rao, C.R. (1973), *Linear Statistical Inferences and its Applications*, Wiley and Sons, 2nd edition, New York.

[21] Rao, C.R. (1988), "Some recent developments in signal detection", *Statistical Decision Theory and related Topics*, Eds. J.O. Berger and S.S. Gupta. Springer and Verlag.

[22] Rao, C.R. and Zhao, L.C. (1993), "Asymptotic behavior of the maximum likelihood estimators of super imposed exponential signals", *IEEE Trans. Signal Proc.* Vol. 41, pp. 1461-1463.

[23] Rao, C.R., Zhao, L.C. and Zhou, B. (1995), "Maximum likelihood estimation of 2-D superimposed exponential signals", *IEEE Trans. Acoust. Speech and Signal Proc.*, Vol. 42, No. 7.

[24] Seber, G.A.F. and Wild, C.J. (1989), *Nonlinear Regression*, John Wiley and Sons, New York.

[25] Walker, A.M. (1971), "On the estimation of a harmonic component in a time series with stationary residuals", *Biometrika*, Vol. 58, No. 1, pp. 21-36.

APPENDIX

Proof of Lemma 1: First we prove the result when $X(m,n)$ is replaced by $e(m,n)$

Consider the following random variables;

$$
\begin{aligned}
Z(m,n) &= e(m,n) \quad \text{if} \quad |X(m,n)| < (mn)^{\frac{3}{4}} \\
&= 0 \qquad\qquad\qquad \text{otherwise}
\end{aligned}
$$

First we will show that $Z(m,n)$ and $e(m,n)$ are equivalent sequences. Consider

$$
\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} P\{e(m,n) \neq Z(m,n)\}
$$

$$
= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} P\{|e(m,n)| > (mn)^{\frac{3}{4}}\}
$$

Now observe that there are at most 2^k k combinations of (m,n)'s such that $mn < 2^k$, therefore we have

$$
\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} P\{|e(m,n)| \geq (mn)^{\frac{3}{4}}\}
$$

$$
\leq \sum_{k=1}^{\infty}\sum_{2^{k-1} \leq r < 2^k} P\{|e(m,n)| \geq r^{\frac{3}{4}}\} \quad \text{[here r=mn]}
$$

$$
\leq \sum_{k=1}^{\infty} k2^k P\{|e(1,1)| \geq 2^{(k-1)\frac{3}{4}}\}
$$

$$
\leq C\sum_{k=1}^{\infty} k2^k \frac{E|e(1,1)|^2}{2^{(k-1)\frac{3}{2}}} \leq C\sum_{k=1}^{\infty} \frac{k}{2^{\frac{k}{2}}} < \infty
$$

Here C is a constant and note that it may represent different constant at different places. Therefore, e(m,n) and Z(m,n) are equivalent sequences. So

$$
P\{e(m,n) \neq Z(m,n) \quad i.o.\} = 0
$$

7

Here i.o. means infinitely often. Let U(m,n)=Z(m,n)-E(Z(m,n)), then

$$\sup_{\alpha,\beta} \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} E(Z(m,n)) \cos(m\alpha)\cos(n\beta) \right|$$

$$\leq \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} |E(Z(m,n))|$$

Since $E(Z(m,n)) \to 0$ as $M,N \to \infty$, therefore as M,N $\to \infty$

$$\frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} |E(Z(m,n))| \to 0$$

Therefore, it is enough to prove that

$$\sup_{\alpha,\beta} \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n) \cos(m\alpha)\cos(n\beta) \right| \overset{a.s.}{\to} 0$$

Now for any fixed $\epsilon > 0$, $-\pi < \alpha, \beta < \pi$ and $0 < h \leq \frac{1}{2(MN)^{\frac{3}{4}}}$, we have

$$P\left\{ \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n) \cos(m\alpha)\cos(n\beta) \right| \geq \epsilon \right\}$$

$$\leq 2e^{-hMN\epsilon} \prod_{m=1}^{M} \prod_{n=1}^{N} E e^{hU(m,n)\cos(m\alpha)\cos(n\beta)}$$

Since $|hU(m,n)\cos(m\alpha)\cos(n\beta)| \leq 1/2$, using $e^x < 1 + x + x^2$ for $|x| < 1/2$, we have

$$2e^{-hMN\epsilon} \prod_{m=1}^{M} \prod_{n=1}^{N} E e^{hU(m,n)\cos(m\alpha)\cos(n\beta)}$$

$$\leq 2e^{-hMN\epsilon}(1 + h^2\sigma^2)^{MN}.$$

Now choose $h = \frac{1}{2(MN)^{\frac{3}{4}}}$, therefore for large M and N

$$P\left\{ \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n) \cos(m\alpha)\cos(n\beta) \right| \geq \epsilon \right\}$$

$$\leq Ce^{-(MN)^{\frac{1}{4}}\epsilon/2} \quad \text{(C is a constant)}.$$

Let $K = M^2N^2$, choose K points, $\theta_1 = (\alpha_1, \beta_1), \ldots,$ $\theta_K = (\alpha_K, \beta_K)$, such that for each point $\theta = (\alpha, \beta) \in (-\pi, \pi)$, we have a point θ_j satisfying

$$|\alpha_j - \alpha| + |\beta_j - \beta| \leq \frac{2\pi}{M^2N^2}$$

Note that

$$\left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n)\{\cos(m\alpha)\cos(n\beta) \right.$$
$$\left. - \cos(m\alpha_j)\cos(n\beta_j)\} \right|$$

$$\leq C\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{MN}{M^2N^2}[m+n] \to 0$$
$$\text{as M,N} \to \infty.$$

Therefore for large M and N, we have

$$P\left\{ \sup_{\alpha,\beta} \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n) \times \right.\right.$$
$$\left.\left. \cos(m\alpha)\cos(n\beta) \right| \geq 2\epsilon \right\}$$

$$\leq P\left\{ \max_{j \leq M^2N^2} \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n) \times \right.\right.$$
$$\left.\left. \cos(m\alpha_j)\cos(n\beta_j) \right| \geq \epsilon \right\}$$

$$\leq CM^2N^2 e^{-(MN)^{\frac{1}{4}}\epsilon/2}.$$

Since $\sum_{t=1}^{\infty} t^2 e^{-t^{\frac{1}{4}}} < \infty$, from Borel Cantelli's lemma, we have

$$\sup_{\alpha,\beta} \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} U(m,n)\cos(m\alpha)\cos(n\beta) \right| \overset{a.s.}{\to} 0$$

Therefore,

$$\sup_{\alpha,\beta} \left| \frac{1}{N}\frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} e(m,n)\cos(m\alpha)\cos(n\beta) \right| \overset{a.s.}{\to} 0$$

Since $P < \infty$, $Q < \infty$ and $|b(i,j)| < \infty$, it proves the lemma.

Table 1
$M = N = 10$

σ	Para.	A_1	A_2	λ_1	λ_2	μ_1	μ_2
.25	LSE	3.998	5.001	2.000	2.500	1.000	1.500
	MSE	1.14E-3	6.30E-4	7.22E-6	2.15E-6	6.60E-6	2.05E-6
	ASV	1.41E-3	5.04E-4	6.04E-6	1.38E-6	6.04E-6	1.38E-6
.50	LSE	4.000	4.998	2.000	2.500	1.000	1.500
	MSE	5.75E-3	2.60E-3	3.12E-5	8.96E-6	3.12E-5	8.05E-6
	ASV	5.64E-3	2.01E-3	2.42E-5	5.53E-6	2.42E-5	5.53E-6
.75	LSE	3.992	4.999	2.000	2.500	1.000	1.500
	MSE	1.25E-2	6.02E-3	6.63E-5	1.87E-5	6.27E-5	1.82E-5
	ASV	1.27E-2	4.53E-3	5.44E-5	1.24E-5	5.44E-5	1.24E-5
1.0	LSE	3.989	5.005	2.000	2.500	1.000	1.500
	MSE	2.15E-2	1.00E-2	1.16E-4	3.55E-5	1.08E-4	3.38E-5
	ASV	2.26E-2	8.06E-3	9.67E-5	2.21E-5	9.67E-5	2.21E-5

Table 2
$M = N = 20$

σ	Para.	A_1	A_2	λ_1	λ_2	μ_1	μ_2
.25	LSE	4.000	5.001	2.000	2.500	1.000	1.500
	MSE	1.24E-4	5.48E-4	3.81E-7	1.11E-7	3.69E-7	1.16E-7
	ASV	3.53E-4	1.26E-4	3.78E-7	8.64E-8	3.78E-7	8.64E-8
.50	LSE	3.999	5.002	2.000	2.500	1.000	1.500
	MSE	5.48E-4	3.38E-4	1.57E-6	3.77E-7	1.65E-6	3.61E-7
	ASV	1.41E-3	5.04E-4	1.51E-6	3.45E-7	1.51E-6	3.45E-7
.75	LSE	3.997	5.004	2.000	2.500	1.000	1.500
	MSE	1.65E-3	8.00E-4	3.15E-6	8.81E-7	2.94E-6	9.91E-7
	ASV	3.17E-3	1.13E-3	3.40E-6	7.77E-7	3.40E-6	7.77E-7
1.0	LSE	4.000	5.005	2.000	2.500	1.000	1.500
	MSE	2.33E-3	1.37E-3	6.82E-6	1.72E-6	6.63E-6	1.71E-6
	ASV	5.64E-3	2.01E-3	6.04E-6	1.38E-6	6.04E-6	1.38E-6

Table 3
$M = N = 30$

σ	Para.	A_1	A_2	λ_1	λ_2	μ_1	μ_2
.25	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	5.37E-5	3.93E-5	8.19E-8	2.15E-8	7.96E-8	1.99E-8
	ASV	1.57E-4	5.60E-5	7.46E-8	1.71E-8	7.46E-8	1.71E-8
.50	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	2.16E-4	1.51E-4	3.09E-7	8.27E-8	2.80E-7	7.79E-8
	ASV	6.27E-4	2.24E-4	2.98E-7	6.82E-8	2.98E-7	6.82E-8
.75	LSE	4.001	5.003	2.000	2.500	1.000	1.500
	MSE	5.31E-4	3.76E-4	6.55E-7	1.75E-7	7.19E-7	1.69E-7
	ASV	1.41E-3	5.04E-4	6.71E-7	1.54E-7	6.71E-7	1.54E-7
1.0	LSE	4.001	5.001	2.000	2.500	1.000	1.500
	MSE	1.04E-3	7.13E-4	1.21E-6	3.22E-7	1.24E-6	2.97E-7
	ASV	2.51E-3	8.95E-4	1.19E-6	2.73E-7	1.19E-6	2.73E-7

Table 4

$$M = N = 40$$

σ	Para.	A_1	A_2	λ_1	λ_2	μ_1	μ_2
.25	LSE	4.001	5.001	2.000	2.500	1.000	1.500
	MSE	1.14E-5	1.60E-5	2.49E-8	5.98E-9	2.37E-8	6.07E-9
	ASV	8.81E-5	3.15E-5	2.36E-8	5.40E-9	2.36E-8	5.40E-9
.50	LSE	4.001	5.000	2.000	2.500	1.000	1.500
	MSE	6.47E-5	7.66E-5	9.57E-8	2.53E-8	9.58E-8	2.43E-8
	ASV	3.53E-4	1.26E-4	9.44E-8	2.16E-8	9.44E-8	2.16E-8
.75	LSE	4.000	5.001	2.000	2.500	1.000	1.500
	MSE	1.20E-4	1.78E-4	2.11E-7	5.69E-8	2.04E-7	5.94E-8
	ASV	7.93E-4	2.83E-4	2.12E-7	4.86E-8	2.12E-7	4.86E-8
1.0	LSE	4.001	5.000	2.000	2.500	1.000	1.500
	MSE	1.74E-4	2.81E-4	3.70E-7	8.61E-8	3.77E-7	9.05E-8
	ASV	1.41E-3	5.04E-4	3.78E-7	8.64E-8	3.78E-7	8.64E-8

Table 5

$$M = N = 50$$

σ	Para.	A_1	A_2	λ_1	λ_2	μ_1	μ_2
.25	LSE	3.999	5.000	2.000	2.500	1.000	1.500
	MSE	2.00E-5	1.37E-5	9.05E-9	2.48E-9	8.65E-9	2.33E-9
	ASV	5.64E-5	2.01E-5	9.67E-9	2.21E-9	9.67E-9	2.21E-9
.50	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	5.89E-5	4.85E-5	3.70E-8	1.02E-8	3.59E-8	9.97E-9
	ASV	2.26E-4	8.06E-5	3.87E-8	8.84E-9	3.87E-8	8.84E-9
.75	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	1.67E-4	1.21E-4	9.29E-8	2.12E-8	9.14E-8	2.01E-8
	ASV	5.08E-4	1.81E-4	8.70E-8	1.99E-8	8.70E-8	1.99E-8
1.0	LSE	3.999	4.999	2.000	2.500	1.000	1.500
	MSE	3.15E-4	2.34E-4	1.80E-7	3.76E-8	1.59E-7	3.89E-8
	ASV	9.02E-4	3.22E-4	1.55E-7	3.54E-8	1.55E-7	3.54E-8